Evaluation of Statistical Estimation Methods for Lognormally Distributed Variables

T. B. PARKIN,* J. J. MEISINGER, S. T. CHESTER, J. L. STARR, AND J. A. ROBINSON

ABSTRACT

Distributions of many chemical, physical, and microbiological properties of soils appear to be lognormal. Several conflicting recommendations exist in the soil science and statistical literature on how to best estimate the population mean, variance, and coefficient of variation of lognormally distributed data. We chose to determine with statistical certainty which of the following three methods is best: (i) the method of moments (method 1); (ii) maximum likelihood (method 2); and (iii) Finney's method (method 3). We assessed the efficacy of these three methods for estimating the mean, variance, and coefficient of variation of lognormal data in the range of sample sizes from $n = 4$ to 100. Three test lognormal populations were used in our evaluation with coefficients of variation that span the range seen for many soil variables (CVs of 50%, 100%, and 200%). We found Finney's method was best for estimating the mean and variance of lognormal data when the coefficient of variation of the underlying lognormal frequency distribution exceeds 100%, below this value the extra computational effort required to implement Finney's technique buys little, relative to the method of moments. Finney's method has not been previously applied by soil scientists, but its superiority over maximum likelihood suggests that the latter should not be generally recommended for estimating the mean, variance and coefficient of variation of lognormal data.

Additional Index Words: **Lognormal, Mean square error, Bias, Efficiency, Soil variables, Monte Carlo simulation.**

THE VARIABILITY of soil properties has received increased interest (Nielsen & Bouma, 1985). The combination of low cost computers and automated analysis systems has enabled scientists to generate large databases for particular soil variables. These large databases have, in turn, allowed the characterization of the variability and frequency distributions of soil variables. Such analyses indicate that the frequency distributions of many physical, chemical, and microbiological soil properties are skewed to the right and are better approximated by the lognormal frequency distribution than by the normal (Gaussian) probability density function (Table 1).

Confusion exists on how to best estimate the mean, variance, and coefficient of variation of lognormally distributed data. This stems from the fact that several statistical procedures for estimating these population parameters have appeared in soil science and statistical literature. (Warrick & Nielsen, 1980; Koch & Link, 1970). A statistically complete evaluation of the most commonly applied methods has not been published in a source accessible to the majority of soil scientists. We undertook the present study to determine the statistical efficacy of three methods (method of moments, maximum likelihood, and Finney's approximation) for estimating the population mean, variance, and coefficient of variation of lognormal data. Of these three, method 2 (maximum likelihood) has been recommended for use by some soil scientists (Warrick & Nielsen, 1980; Folurenso & Rolston, 1984; Parkin et al. 1985), although, as we will show, it is generally inferior to Finney's method (method 3) as well as the more commonly applied method of moments (method 1). Finney's method for estimating the mean, variance, and coefficient of variation of lognormal data has only rarely been applied to soils data. (White et al., 1987; Parkin, 1987; Parkin et al. 1987).

The present study is a theoretical one in the sense that we investigated the properties of the three estimation methods in a "world" where the answers were known. Description of any new statistical estimation method often incorporates an evaluation of the method for a family of known probability density functions that cover the range of distributions seen or expected

T.B. Parkin, J.J. Meisinger, and J.L. Starr, USDA-ARS, BARC. Soil Nitrogen and Environmental Chemistry Lab., Beltsville, MD 20705; S.T. Chester and J.A. Robinson, The Upjohn Company, Kalamazoo, MI 49001. Contribution of the USDA-ARS and The Upjohn Company. Received 29 May 1987. *Corresponding author.

Published in Soil Sci. Soc. Am. J. 52:323–329 (1988).

Table 1. Abbreviated survey of soil variables which have been reported to be approximately lognormally distributed.

Soil variable	n	% CV	Method†	Reference
		Structure		
Aggregate size	36	119–182‡	2	Allmaras et al., 1965
		Water		
Water flux	81–100	31–499	1	Jones & Wagenet, 1984
Hydraul. conduct.	20	102–132	1	Warrick et al., 1977
Hydraul. conduct	20	157–321	2	Warrick et al., 1977
		N		
Nitrate	20	21–87	1	Cameron & Wild, 1984
	12–30	27–83	2	Ruess, et al., 1977
	49	29	2	Tabor, et al., 1985
	149	51–56	2	White et al., 1987
d ^{15}N	46	11.1	1	Sells et al., 1986
Total N	46	22.8	1	Sells et al., 1986
		Gases		
CO_2, pore space	137–195	71–105‡	2	Focht et al., 1979
N_2O, pore space	135–205	135–277‡	2	Focht et al., 1979
N_2O, surface flux	36	282–379	2	Folorunso & Rolston, 1984
		Denitrification		
Surface chambers	36	161–508	2	Folorunso & Rolston, 1984
	3	1–115	1	Duxbury & McConnaughey, 1986
Intact soil cores	4	44–120	1	Burton & Beauchamp, 1984
	4–50	95–250	2	Parkin et al., 1985
	36	128–383	3	Parkin et al., 1987
Enzyme activity	36	18.7–58.4	3	Parkin et al., 1987
		Bacterial numbers		
Plant leaves	24	15.3–236‡	2	Hirano et al., 1982
Rhizosphere	40–60	26.7–157‡	2	Loper et al., 1984

† Method refers to the statistical method used for calculating estimates of the mean, variance, and coefficient of variation.
‡ Coefficients of variation were not reported but were calculated from the variance of the log transformed values using method 2 (Eq. [6]) of this paper.

for real data. We followed the same approach in the study described in this paper. We evaluated the ability of three methods to estimate the population mean, variance, and coefficient of variation of lognormal data given three lognormal populations where the mean and variances were known. This was done for sample sizes of $n = 4$ to 100, making our results relevant to most typical studies in soil sciences. The lognormal distributions we used to evaluate the three estimation methods are representative of lognormal distributions that describe the frequency distributions of several soil properties (Table 1).

THEORY

Three different methods of estimation were evaluated with respect to how well the population mean, variance, and coefficient of variation were estimated from sample data. These methods are: (i) the method of moments, (ii) the method of maximum likelihood, and (iii) the uniformly minimum variance unbiased estimator (UMVUE) method of Finney (1941). The implementation of these methods in estimating the population mean (α), variance (β^2), and coefficient of variation (CV) is presented below.

Method 1

This is the method of moments which is typically applied when a normal frequency distribution is observed or when the sample size is too small to adequately determine the true population frequency distribution but a normal distribution is assumed. The variance estimate which appears in Eq. [2] is not strictly a method of moments as the term $n - 1$ rather than n appears in the denominator. The term $n - 1$ was used here because of the known bias which exists for low sample numbers when n is used (Snedecor and Cochran, 1967). This method was included because it is the most commonly used method in soil science and therefore provides a well known benchmark. The mean, variance, and CV are calculated using untransformed data as described by Eq. [1, 2, and 3], respectively.

$$m_1 = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad [1]$$

$$s_1^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - m_1^2) \qquad [2]$$

$$cv_1 = s_1/m_1 \qquad [3]$$

where
x_i = the untransformed ith observation,
n = the number of observations,
m_1 = the estimate of the population mean (α),
s_1^2 = the estimate of the population variance, (β^2), and
cv_1 = the estimate of the coefficient of variation, CV. Subscripts refer to method 1.

Method 2

The maximum likelihood method has been recommended when the observations are thought to be lognormally distributed (Warrick & Nielsen, 1980). Estimates of α, β^2, and CV are given by (Aitchinson & Brown, 1957) and are presented in Eq. [4], [5], and [7], respectively:

$$m_2 = \exp(\bar{\mu} + \sigma^2/2) \qquad [4]$$

$$s_2^2 = m_2^2\,[\exp(\sigma^2) - 1] \qquad [5]$$

$$cv_2 = s_2/m_2 = [\exp(\sigma^2) - 1]^{1/2} \qquad [6]$$

where

$$\bar{\mu} = \frac{1}{n}\sum_{i=1}^{n}\ln(x_i) \qquad [7]$$

and

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\ln(x_i) - \bar{\mu}^2)\,. \qquad [8]$$

Method 3

These uniformly minimum variance unbiased estimators (UMVUE) were developed independently by Finney (1941) and Sichel (1952) and have been typically applied to the analysis of geological data (Krige, 1981; Koch & Link, 1970). Estimators of α, β^2, and CV are given by Eq. [9], [10], and [11], respectively:

$$m_3 = \exp(\bar{\mu})\,\Psi\,(\sigma^2/2) \qquad [9]$$

$$s_3^2 = \exp(2\bar{\mu})\left\{\Psi\,(2\,\sigma^2) - \Psi\left[\frac{(n-2)}{(n-1)}\,\sigma^2\right]\right\} \qquad [10]$$

$$cv_3 = s_3/m_3 \qquad [11]$$

where Ψ = the power series given in Eq. [12].

$$\Psi\,(t) = 1 + \frac{t(n-1)}{n} + \frac{t^2\,(n-1)^3}{n^2(n+1)2!} + \frac{t^3(n-1)^5}{n^3(n+1)\,(n+3)3!}$$

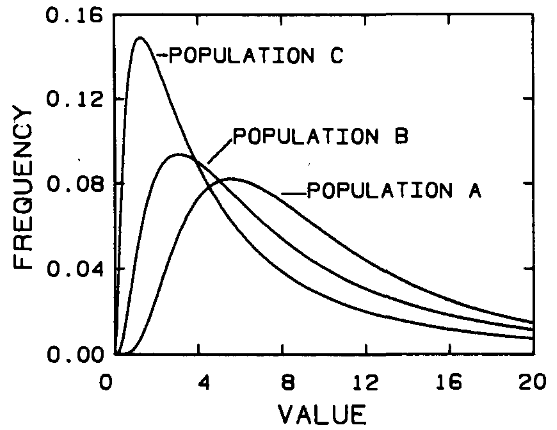

Fig. 1. Probability density functions of the three lognormal populations evaluated in this study. The mean of each population (in untransformed units) is 10 and variances of the populations are 25 (pop. A), 100 (pop. B), and 400 (pop. C), respectively.

$$+ \frac{t^4(n-1)^7}{n^4(n+1)(n+3)(n+5)4!} + \dots \quad [12]$$

In this study Eq. [12] was evaluated until the final term accounted for <1% of the sum of the preceeding terms. This usually required the calculation of 6 to 10 terms.

EVALUATION OF THE METHODS

Three different lognormal populations with known parameters were used in these evaluations. Figure 1 shows the probability density functions of the three populations. Population A has a CV of 50% and is only slightly skewed. Population C is highly skewed (CV = 200%) and population B is intermediate with respect to skewness (CV = 100%). These populations represent a range of skewness often observed for soil variables (Table 1). Each population had a mean of 10 and variances of 25, 100, and 400, respectively. Other characteristics of the test populations are presented in Table 2.

For each population parameter estimated, the three estimators were compared with respect to their relative efficiency by comparing their mean square errors (Barnett, 1973). The mean square error (MSE) is the average squared deviation of an estimate from the parameter it is estimating. It is a combination of both the estimator's variance (average squared deviation about the estimator's expected value) and bias (the systematic under- or over-estimation of the parameter being estimated, Eq. [13])

$$\text{MSE} = \text{variance} + \text{bias}^2 . \quad [13]$$

The use of MSE allows one to compare both biased and unbiased estimators. When an estimator is unbiased, the MSE is equal to the variance or variability of the estimator. Thus, the estimator with the smallest MSE is the one that has the smallest (squared) deviation from the parameter being estimated regardless of whether it is biased or not. The formulas for the variance and bias of each of the estimators of the lognormal mean and variance are presented in the appendix.

Monte Carlo Simulations

Exact solutions for the MSE of the UMVUE variance estimator and for the CV estimators do not exist, thus, these parameters were evaluated empirically using Monte Carlo simulation techniques. The influence of sample size (n) on these estimators was determined by selecting from 4 to 100 (incremented by 2) observations from each population and calculating the UMVUE variance estimator and CVs using

Table 2. Statistical properties of 3 populations used in the Monte Carlo simulations.

Population	Mean (α)	Median	Mode	Skewness	Variance (β^2)	Mean of logs (μ)	SD of logs (σ)
A	10.0	8.944	7.155	1.625	25.0	2.191	0.4724
B	10.0	7.071	3.536	4.000	100.0	1.956	0.8326
C	10.0	4.472	0.894	14.000	400.0	1.498	1.2690

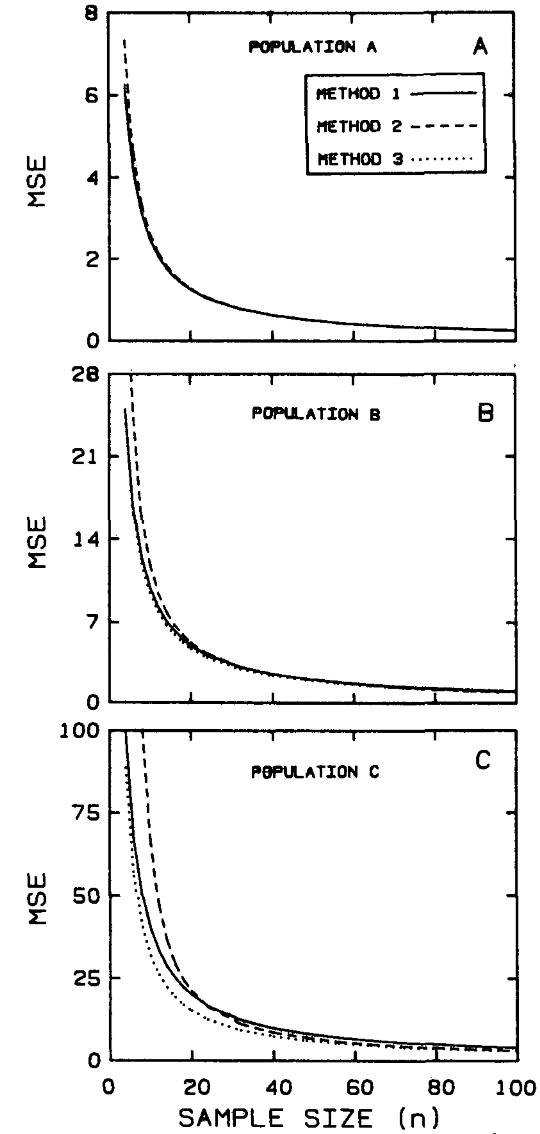

Fig. 2. Influence of sample size on mean square error of the three statistical methods used to estimate the mean of populations A (Fig. 2A), B (Fig. 2B), and C (Fig. 2C). Presented are the exact solutions of the mean square errors.

all three methods. Ten thousand Monte Carlo simulations were run for each sample size, using an algorithm for generating random variates from a lognormal distribution. The bias, variance, and mean square error of each estimator was calculated from Eq. [14], [15], and [16], respectively.

$$\text{BIAS}(P) = \frac{1}{10\,000} \sum_{k=1}^{10000} (P_k - P) \quad [14]$$

$$\text{VAR}(P) = \frac{1}{10\,000} \sum_{k=1}^{10000} (P_k - \bar{P})^2 \quad [15]$$

$$\text{MSE}(P) = \text{VAR}(P) + \text{BIAS}(P)^2 \quad [16]$$

where
10 000 = the number of Monte Carlo simulations,
P = the true population parameter value,
P_k = the estimate of the population parameter obtained from the kth simulation, and
$\bar{P}$ = the average parameter estimate of the 10 000 simulations.

RESULTS AND DISCUSSION

Estimating the Mean

For populations A and B (Fig. 2A, B) all three methods were nearly the same with respect to the behavior of their mean square errors (MSEs) in estimating the mean. The MSEs of methods 1 and 3 were indistinguishable over all sample sizes while the MSE of method 2 was slightly higher than the other methods for small sample sizes ($n < 20$). At higher sample numbers ($n > 40$) the MSEs of all three methods were essentially equal.

With population C (Fig. 2C) a difference in the MSE curves could be discerned. Over all sample sizes method 3 had a lower MSE than the other methods. For sample sizes of $n < 20$ method 2 exhibited larger MSEs than either methods 1 or 3, however, for $n > 20$ method 2 had lower MSEs than method 1, and the MSE of method 2 approached levels exhibited by method 3.

Estimators of the mean given by methods 1 and 3 are unbiased, however, the method 2 estimator yields biased values of the mean (Table 3). The bias decreases with increasing sample size, however, with increasing population skewness bias increases for a given sample size. Thus, for sample sizes of $n = 4$, method 2 overestimates the mean by 3.28%, 14.46%, and 73.29%, for populations A, B, and C, respectively.

There have been few studies which evaluate various statistical methods as applied to skewed data. Evaluation of procedures for estimating mean and standard deviation of a three-parameter lognormal distribution was conducted by Cohen and Whitten (1980), however, the influence of sample size was not considered. Warrick and Nielsen (1980) compared estimators of the mean using methods 1 and 2 (Eq. [1] and [4], respectively) using a single data set containing 20 observations and concluded that method (i) yields underestimates of the population mean. These investigators were only working with a single small data set and α was unknown. Results presented here indicate that method 1 is unbiased and does not underestimate the mean. Rather method 2 is biased and yields overestimates. The results of this study support the theoretical work of Finney (1941) and Sichel (1952) which predicted that application of Eq. [4] to samples with $n < 100$ would yield over-estimates of α and that method 3 (Eq. [8]) is the minimum variance unbiased estimator.

Table 3. Influence of sample size and population skewness on bias of the method 2 mean estimator.

	% bias of the mean population		
Sample size (n)	A	B	c
4	3.28	14.46	73.29
12	1.05	4.11	14.15
20	0.63	2.41	7.94
40	0.31	1.19	3.79
60	0.21	0.79	2.49
100	0.12	0.47	1.48

Estimating the Variance

The MSE of the three variance estimators for the test populations are shown in Fig. 3. For each population, method 3 exhibited the smallest MSEs throughout the entire range of sample sizes. For population A (Fig. 3A) the MSE of each variance estimator was nearly the same for small sample sizes ($n < 15$), however, for larger sample sizes, method 1 had higher MSEs than methods 1 or 2. As population skewness increased, method 3 had the lowest MSEs over the entire range of sample sizes (Fig. 3B, C). With all populations, for low sample numbers, method 2 exhibited high MSEs and for larger samples the performance of method 2 approached that of method 3.

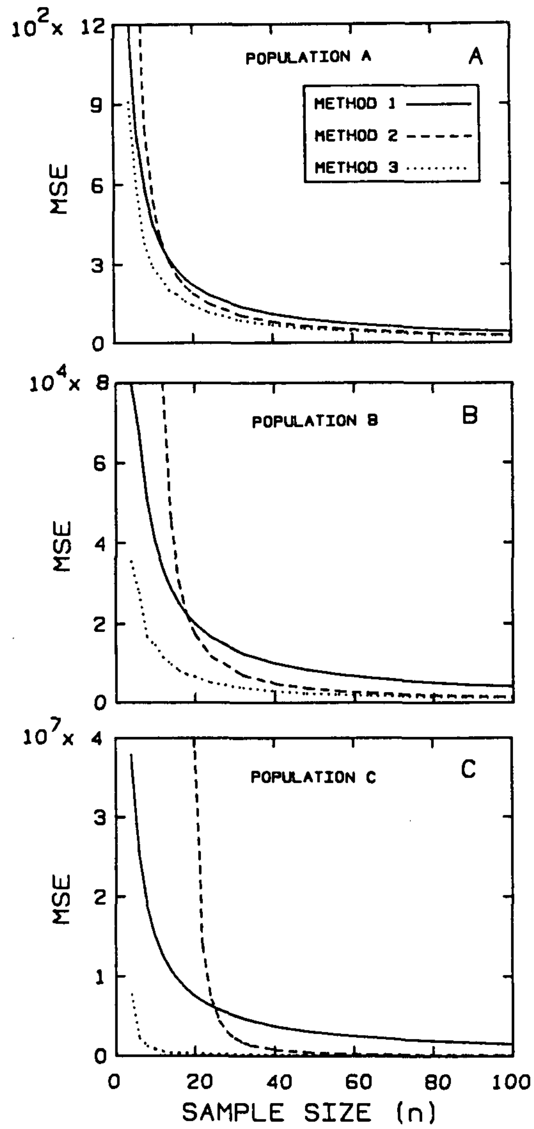

Fig. 3. Influence of sample size on mean square error of the three statistical methods used to estimate the variance of populations A (Fig. 3A), B (Fig. 3B), and C (Fig. 3C). Exact solutions of the mean square errors for methods 1 and 2 are presented. Mean square errors for method 3 were estimated using Monte Carlo simulation.

Method 1 performed better than method 2 for $n < 20$ but the MSEs of method 1 were higher than those of methods 2 or 3 at larger sample sizes.

The variance estimators of methods 1 and 3 were unbiased, however, method 2 yielded biased estimates of the population variance (Table 4). Even for populations having low skewness (population A), for a sample size of 4, method 2 yielded estimates which overestimated the population variance by 52%. The bias of this estimator increased with increasing population skewness, for any given sample size.

Estimating the Coefficient of Variation

Although the CV has no function in statistical testing it is commonly used to indicate general differences in relative dispersion between variables. Also, the CV may be a useful index for deciding which statistical methods to use in estimating the mean and variance. Because of the widespread use of the CV statistic, an evaluation of different CV estimators was included in this study.

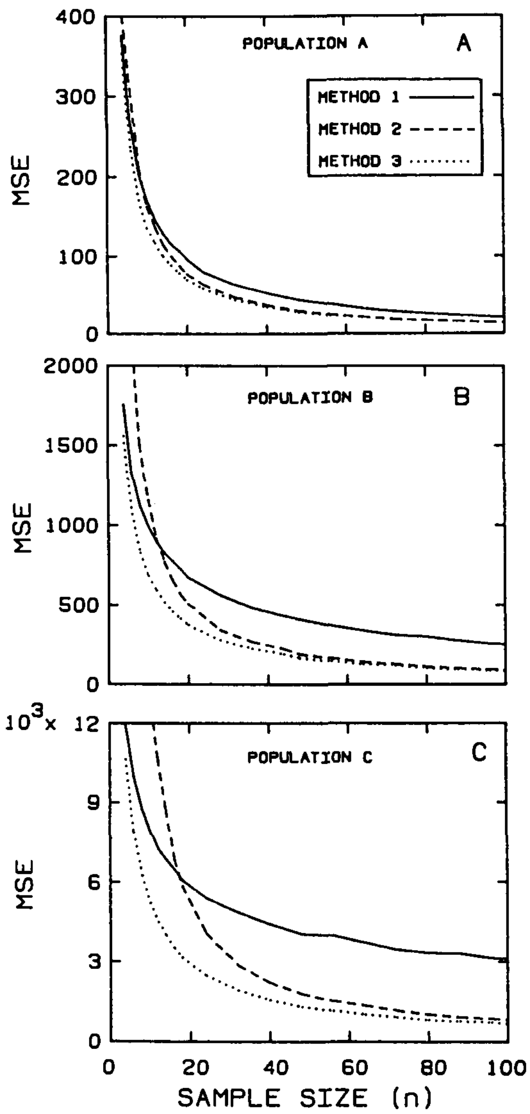

Fig. 4. Influence of sample size on mean square error of the three statistical methods used to estimate the coefficient of variation of populations A (Fig. 4A), B (Fig. 4B), and C (Fig. 4C). The mean square errors for each sample size were estimated using Monte Carlo simulation.

Method 3 yielded lower MSEs than the other methods over the entire range of sample sizes for the three test populations (Fig. 4). The clear superiority of the method 3 estimator in calculating CV is demonstrated with the highly skewed population (Fig. 4C).

Method 2 performed poorly at low sample numbers and had higher MSEs than either method 1 or method 3. With increasing sample numbers, the MSEs of method 2 decreased to levels below those of method 1 and approached the MSEs exhibited by method 3.

CONCLUSIONS AND RECOMMENDATIONS

Historically, there has been much interest in lognormal distributions, as many natural variables are skewed and can be approximated by this class of frequency distributions. Consequently, there have been several studies of parameter estimation methods for the lognormal distribution (Aitchison & Brown, 1957; Koch & Link, 1970; Cohen & Whitten, 1980; Vevjevich & Obeysekera, 1984; Rukhin, 1986).

However, we found the existing literature to be inconclusive or not easily utilizable for practical applications in soil science. This in no way diminishes the importance of the existing literature because some of these studies were conducted before inexpensive computing time was available, which precluded the evaluation of methods for which theory was insufficient. In other cases the work done in the statistical sciences simply did not find its way into soil science. This study was conducted to fill knowledge gaps in the literature and to partially bridge the communication gap between statistics and soil science.

With the above objective in mind, presented in Table 5 are summary recommendations of the three

Table 4. Influence of sample size and population skewness on bias of the method 2 variance estimator.

	% bias of the variance population		
Sample size (n)	A	B	C
4	52.11	3110	ud†
12	11.88	59.51	652.5
20	6.73	29.74	164.9
40	3.23	13.21	53.6
60	2.13	8.49	31.8
100	1.26	4.95	17.5

† Undefined at given sample number and population variance.

Table 5. Summary recommendations of estimation methods for three sampling intensity ranges for samples from three lognormal populations.

Lognormal population	Sample size (n)	Recommended method†		
		Mean	Variance	CV
A	4–20	1,3	1,3	3
	20–40	1,3	3	3,2
	40–100	1,3,2	3,2	3,2
B	4–20	3,1	3	3
	20–40	3,1	3	3
	40–100	3,1,2	3	3,2
C	4–20	3	3	3
	20–40	3	3	3
	40–100	3,2	3	3,2

† When more than one method is recommended, the methods are presented in order of most to least preferable.

methods evaluated in this study. These recommendations are based, primarily, on the relative magnitudes of the MSEs of the estimators for the given populations and sample ranges. When no substantial differences in MSEs were apparent, two secondary criteria (i) bias of the estimator, and (ii) ease of implementation of the method, were used to rank the methods. In these cases when more than one method is recommended the recommendations are made in order of decreasing preference. As indicated in Table 5, the performance of method 3 was equal to or better than the other methods in estimating the mean, variance, and CV over all sample size ranges for all three lognormal populations, however for slightly skewed populations method 1 performed nearly as well and is much easier to implement. For highly skewed populations method 3 was clearly superior than methods 1 or 2. We evaluated the three methods for an additional lognormal population having a CV of 500% (data not shown) and found results similar to those obtained for population C.

It should be cautioned that the results of these analyses are strictly valid only for samples drawn from true lognormal distributions. In sampling we rarely have the luxury of knowing the true distribution function of the underlying population, therefore certain assumptions are required in order to summarize data. However, if the data are skewed it is logical to use a skewed model, such as the lognormal, rather than a symmetric model. We are currently working to evaluate the robustness of these statistical methods to deviations from strict lognormal populations.

APPENDIX

Presented below are equations for the variance and bias of the mean and variance estimators using methods 1, 2, and 3. These quantities were used to calculate the mean square error of the estimators as described in Eq. [13].

Definitions

Let $X_1, \ldots, X_n$ be normally distributed with mean $= \mu$ and variance $= \sigma^2$, then if $Y_i = e^{X_i}$; $Y_1, \ldots, Y_n$ is lognormally distributed with mean $= \alpha$ and variance of β^2. The expected value and variance of Y [$E(Y)$ and VAR(Y), respectively] are given by Eq. [14] and [15]:

$$E(Y) = \alpha = \exp(\mu + \sigma^2/2) \quad [14]$$

$$\mathrm{VAR}(Y) = \beta^2 = \alpha^2[\exp(\sigma^2) - 1]. \quad [15]$$

Variance and Bias of the Mean

Presented in Eq. [16] through [21] are the variance and bias of the mean estimators of methods 1, 2, and 3 (m_1, m_2, and m_3 calculated by Eq. [1], [4], and [9], respectively):

$$\mathrm{VAR}(m_1) = \frac{\alpha^2}{n}[\exp(\sigma^2) - 1] \quad [16]$$

$$\mathrm{BIAS}(m_1) = 0 \quad [17]$$

$$\mathrm{VAR}(m_2) = \exp[2\mu + (\sigma^2/n)]\,\{\exp(\sigma^2/n)$$
$$[1 - 2\sigma^2/(n-1)]^{\frac{-(n-1)}{2}} - [1 - \sigma^2/(n-1)]^{-(n-1)}\} \quad [18]$$

$$\mathrm{BIAS}(m_2) = \alpha\,\{\exp\left[\frac{-\sigma^2(n-1)}{2n}\right]$$
$$\left[1 - \sigma^2/(n-1)\right]^{\frac{-(n-1)}{2}} - 1\} \quad [19]$$

$$\mathrm{VAR}(m_3) = \alpha^2\,\{\exp(\sigma^2/n)\,\Psi\left(\frac{\sigma^4(n-1)}{2n^2}\right) - 1\} \quad [20]$$

$$\mathrm{BIAS}(m_3) = 0. \quad [21]$$

Variance and Bias of the Variance

Presented in Eq. [22] through [27] are the variance and bias of the variance estimators of methods 1, 2, and 3 (s_1^2, s_2^2, and s_3^2 calculated by Eq. [2], [5], and [10], respectively):

$$\mathrm{VAR}(s_1^2) = \frac{\alpha^4}{n}\{\exp(6\sigma^2) - 4\exp(3\sigma^2) - \exp(2\sigma^2) + 8\exp(\sigma^2) - 4\} + \frac{2\alpha^4}{n(n-1)}\{\exp(2\sigma^2) - 2\exp(\sigma^2) + 1\} \quad [22]$$

$$\mathrm{BIAS}(s_1^2) = 0 \quad [23]$$

$$\mathrm{VAR}(s_2^2) = \exp[4\mu + (8/n)\sigma^2]\{[1 - 8\sigma^2/(n-1)]^{\frac{-(n-1)}{2}} - 2\,[1 - 6\sigma^2/(n-1)]^{\frac{-(n-1)}{2}} + [1 - 4\sigma^2/(n-1)]^{\frac{-(n-1)}{2}}\}$$
$$- \exp[4\mu + (4/n)\sigma^2]\,\{(1 - 4\sigma^2/(n-1)]^{\frac{-(n-1)}{2}} - [1 - 2\sigma^2/(n-1)]^{\frac{-(n-1)}{2}}\}^2 \quad [24]$$

$$\mathrm{BIAS}(s_2^2) = \alpha^2\,\{\exp[-(n-2)\sigma^2/n]\,\{[1 - 4\sigma^2/(n-1)]^{\frac{-(n-1)}{2}} - [1 - 2\sigma^2/(n-1)]^{\frac{-(n-1)}{2}}\} - \exp(\sigma^2) + 1\} \quad [25]$$

$$\mathrm{VAR}(s_3^2) = \text{no closed form solution} \quad [26]$$

$$\mathrm{BIAS}(s_3^2) = 0. \quad [27]$$

The equations presented above were solved over a range of sample sizes of $n = 4$ to 100 (incremented by 2) for the three lognormal populations described in Table 2. A closed form solution of equation 26 does not exist, hence the variance of s_3^2 was calculated using the Monte Carlo simulation techniques described previously.

REFERENCES

Allmaras, R.R., R.E. Burwell, W.B. Voorhees, and W.E. Larson. 1965. Aggregate size distribution in the row zone of tillage experiments. Soil Sci. Soc. Am. Proc. 29:645–650.

Aitchison, J., and J.A.C. Brown. 1957. The lognormal distribution. Cambridge Univ. Press., London.

Barnett, V. 1973. Comparative statistical inference. John Wiley & Sons Inc., New York. p. 122.

Burton, D.L., and E.C. Beauchamp. 1984. Field techniques using the acetylene blockage of nitrous oxide reduction to measure denitrification. Can. J. Soil Sci. 64:555–562.

Cameron, K.C., and A. Wild. 1984. Potential aquifer pollution leaching following the plowing of temporary grassland. J. Environ. Qual. 13:274–277.

Cohen, A.C., and B.J. Whitten. 1980. Estimation in the three-parameter lognormal distribution. J. Am. Stat. Assoc. 75:399–404.

Duxbury, J.M., and P.K. McConnaughey. 1986. Effect of fertilizer source on denitrification and nitrous oxide emissions in a maize-field. Soil Sci. Soc. Am. J. 50:644–648.
Finney, D.J. 1941. On the distribution of a variate whose logarithm is normally distributed. Royal Stat. Soc. London J. Supplement 7:155–161.
Focht, D.C., L.H. Stolzy, and B.D. Meek. 1979. Sequential reduction of nitrous oxide under field conditions as brought about by organic amendments and irrigation management. Soil Biol. Biochem. 11:37–46.
Folorunso, O.A., and D.E. Rolston. 1984. Spatial variability of field-measured denitrification gas flux. Soil Sci. Soc. Am. J. 48:1214–1219.
Hirano, S.S., E.V. Nordheim, D.C. Arny, and C.D. Upper. 1982. Lognormal distribution of epiphytic bacterial populations on leaf surfaces. Appl. Environ. Microbiol. 44:695–700.
Jones, A.J., and R.J. Wagenet. 1984. In situ estimation of hydraulic conductivity using simplified methods. Water Resour. Res. 20:1620–1626.
Koch, George S., and Richard F. Link. 1970. Statistical analysis of geological data. John Wiley & Sons, Inc., New York.
Krige, D.G. 1981. Lognormal-de wijsian geostatistics for ore evaluation. South African Inst. of Mining and Metallurgy, Johannesburg, South Africa.
Loper, J.E., T.V. Suslow, and N.M. Schroth. 1984. Lognormal distribution of bacterial populations in the rhizosphere. Phytopathology 74:1454–1460.
Nielsen, D.R., and J. Bouma (ed.) 1985. Soil spatial variability. Proc. Workshop of the ISSI and SSSA, Las Vegas, NV 30 Nov. 1984-1 Dec. 1984. PUDOC, Wageningen, The Netherlands.
Parkin, T.B., A.J. Sexstone, and J.M. Tiedje. 1985. Comparison of field denitrification rates determined by acetylene-based soil core and nitrogen-15 methods. Soil Sci. Soc. Am. J. 49:94–99.
Parkin, T.B. 1987. Soil microsites as a source of denitrification variability. Soil Sci. Soc. Am. J. 51:1194–1199.
Parkin, T.B., J.L. Starr, and J.J. Meisinger. 1987. The influence of sample size on measurement of soil denitrification. Soil Sci. Soc. Am. J. 51:1492–1501.
Rukhin, A.L. 1986. Improved estimation in lognormal models. J. Am. Stat. Assoc. 81:1046–1049.
Ruess, J.O., P.N. Soltanpour, and A.E. Ludwick. 1977. Sampling distribution of nitrates in irrigated fields. Agron. J. 69:588–592.
Sells, F., R.E. Karamanos, and R.G. Kachanoski. 1986. The spatial variability of nitrogen-15 and its relation to the variability of other soil properties. Soil Sci. Soc. Am. J. 50:105–110.
Sichel, H.S. 1952. New methods in the statistical evaluation of mine sampling. Trans. Inst. Min. Metall. (London) Sec. B. 61:261–288.
Snedecor, G.W., and W.G. Cochran. 1967. Statistical methods. 6th ed., Iowa State Univ., Ames, IA.
Tabor, J.A., A.W. Warrick, D.E. Myers, and D.A. Pennington. 1985. Spatial variability of nitrate in irrigated cotton: II. Soil nitrate and correlated variables. Soil. Sci. Soc. Am. J. 49:390–394.
Vevjevich, V., and Obeysekera, J.T.B. 1984. Estimation of skewness of hydrologic variables. Water. Resour. Res. 20:935–943.
Warrick, A.W., G.J. Mullen, and D.R. Nielsen. 1977. Scaling field measured soil hydraulic properties using a similar media concept. Water Resour. Res. 13:355–362.
Warrick, A.W., and D.R. Nielsen. 1980. Spatial variability of soil physical properties in the field. p. 385. *In* D. Hillel (ed.) Applications of soil physics. Academic Press, New York.
White, R.E., R.A. Haigh, and J.H. Mcduff. 1987. Frequency distributions and spatially dependent variability of ammonium and nitrate concentrations in soil under grazed and ungrazed grasslands. Fert. Res. 11:193–208.

ERRATA

Evaluation of Statistical Estimation Methods for Lognormally Distributed Variables

T.B. PARKIN, J.J. MEISINGER, S.T. CHESTER, J.L. STARR, AND J.A. ROBINSON

Soil Sci. Soc. Am. J. 52:323–329. (Mar.–Apr. 1988)

The following corrections should be made:

1. Sentence 7 of the abstract should read

Finney's method has been rarely applied by soil scientists, but its superiority over the maximum likelihood method suggests that the latter should not be generally recommended for estimating the mean, variance, and coefficient of variation of lognormal data.

2. In the body of Table 1, line 11, column 4 should read 1,2,3 instead of 2.
3. In Eq. [2] the term $(x_i - m_l^2)$ should be $(x_i - m_l)^2$.
4. In Eq. [8] the term $(\ln(x_i) - \bar{\mu}^2)$ should be $(\ln(x_i) - \bar{\mu})^2$.